

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/479,862

DATE: 12/20/2000
TIME: 22:20:58

INPUT SET: S36235.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT: Takanori OKURA
6 Kakuji TORIGOE
7 Masahi KURIMOTO
8
9 (ii) TITLE OF INVENTION: GENOMIC DNA ENCODING A POLYPEPTIDE CAPABLE
10 OF INDUCING THE PRODUCTION OF INTERFERON-
11
12 (iii) NUMBER OF SEQUENCES: 35
13
14 (iv) CORRESPONDENCE ADDRESS:
15 (A) ADDRESSEE: BROWDY AND NEIMARK
16 (B) STREET: 419 Seventh Street, N.W., Suite 300
17 (C) CITY: Washington
18 (D) STATE: D.C.
19 (E) COUNTRY: USA
20 (F) ZIP: 20004
21
22 (v) COMPUTER READABLE FORM:
23 (A) MEDIUM TYPE: Floppy disk
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: Patent In Release #1.0, Version #1.30
27
28 (vi) CURRENT APPLICATION DATA:
29 (A) APPLICATION NUMBER: 09/479,862
30 (B) FILING DATE:
31 (C) CLASSIFICATION:
32
33 (vii) PRIOR APPLICATION DATA:
34 (A) APPLICATION NUMBER: 08/884,324
35 (B) FILING DATE:
36
37 (viii) ATTORNEY/AGENT INFORMATION:
38 (A) NAME: BROWDY, Roger L.
39 (B) REGISTRATION NUMBER: 25,618
40 (C) REFERENCE/DOCKET NUMBER: OKURA=1
41
42 (ix) TELECOMMUNICATION INFORMATION:
43 (A) TELEPHONE: 202-628-5197
44 (B) TELEFAX: 202-737-3528
45
46 (2) INFORMATION FOR SEQ ID NO: 1:

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/479,862DATE: 12/20/2000
TIME: 22:20:58

INPUT SET: S36235.raw

47
48 (i) SEQUENCE CHARACTERISTICS:
49 (A) LENGTH: 157 amino acids
50 (B) TYPE: amino acid
51 (D) TOPOLOGY: linear
52
53 (ii) MOLECULE TYPE: peptide
54
55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
56
57 Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn
58 1 5 10 15
59 Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp
60 20 25 30
61 Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile
62 35 40 45
63 Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile
64 50 55 60
65 Ser Val Lys Cys Glu Lys Ile Ser Xaa Leu Ser Cys Glu Asn Lys Ile
66 65 70 75 80
67 Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys
68 85 90 95
69 Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys
70 100 105 110
71 Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu
72 115 120 125
73 Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu
74 130 135 140
75 Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp
76 145 150 155
77
78 (2) INFORMATION FOR SEQ ID NO: 2:
79
80 (i) SEQUENCE CHARACTERISTICS:
81 (A) LENGTH: 1120 base pairs
82 (B) TYPE: nucleic acid
83 (C) STRANDEDNESS: double
84 (D) TOPOLOGY: linear
85
86 (ii) MOLECULE TYPE: cDNA to mRNA
87
88 (iii) HYPOTHETICAL: No
89
90 (iv) ANTI-SENSE: No
91
92 (vi) ORIGINAL SOURCE:
93 (A) ORGANISM: human
94 (F) TISSUE TYPE: liver
95
96 (ix) FEATURE:
97 (A) NAME/KEY: 5'UTR
98 (B) LOCATION: 1..177
99 (C) IDENTIFICATION METHOD: E

RAW SEQUENCE LISTING PATENT APPLICATION US/09/479,862

DATE: 12/20/2000
TIME: 22:20:58

INPUT SET: S36235.raw

```

100      (A) NAME/KEY: leader peptide
101      (B) LOCATION: 178..285
102      (C) IDENTIFICATION METHOD: S
103      (A) NAME/KEY: mat peptide
104      (B) LOCATION: 286..756
105      (C) IDENTIFICATION METHOD: S
106      (A) NAME/KEY: 3'UTR
107      (B) LOCATION: 757..1120
108      (C) IDENTIFICATION METHOD: E
109
110      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
111
112      GCCTGGACAG TCAGCAAGGA ATTGTCTCCC AGTGCATTTT GCCCTCCTGG CTGCCAACTC      60
113      TGGCTGCTAA AGCGGCTGCC ACCTGCTGCA GTCTACACAG CTTCGGGAAG AGGAAAGGAA      120
114      CCTCAGACCT TCCAGATCGC TTCCTCTCGC AACAAACTAT TTGTCGCAGG AATAAAG      177
115      ATG GCT GCT GAA CCA GTA GAA GAC AAT TGC ATC AAC TTT GTG GCA ATG      225
116      Met Ala Ala Glu Pro Val Glu Asp Asn Cys Ile Asn Phe Val Ala Met
117      -35 -30 -25
118      AAA TTT ATT GAC AAT ACG CTT TAC TTT ATA GCT GAA GAT GAT GAA AAC      273
119      Lys Phe Ile Asp Asn Thr Leu Tyr Phe Ile Ala Glu Asp Asp Glu Asn
120      -20 -15 -10 -5
121      CTG GAA TCA GAT TAC TTT GGC AAG CTT GAA TCT AAA TTA TCA GTC ATA      321
122      Leu Glu Ser Asp Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile
123      1 5 10
124      AGA AAT TTG AAT GAC CAA GTT CTC TTC ATT GAC CAA GGA AAT CGG CCT      369
125      Arg Asn Leu Asn Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro
126      15 20 25
127      CTA TTT GAA GAT ATG ACT GAT TCT GAC TGT AGA GAT AAT GCA CCC CGG      417
128      Leu Phe Glu Asp Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg
129      30 35 40
130      ACC ATA TTT ATT ATA AGT ATG TAT AAA GAT AGC CAG CCT AGA GGT ATG      465
131      Thr Ile Phe Ile Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met
132      45 50 55 60
133      GCT GTA ACT ATC TCT GTG AAG TGT GAG AAA ATT TCA AYT CTC TCC TGT      513
134      Ala Val Thr Ile Ser Val Lys Cys Glu Lys Ile Ser Xaa Leu Ser Cys
135      65 70 75
136      GAG AAC AAA ATT ATT TCC TTT AAG GAA ATG AAT CCT CCT GAT AAC ATC      561
137      Glu Asn Lys Ile Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile
138      80 85 90
139      AAG GAT ACA AAA AGT GAC ATC ATA TTC TTT CAG AGA AGT GTC CCA GGA      609
140      Lys Asp Thr Lys Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly
141      95 100 105
142
143
144      CAT GAT AAT AAG ATG CAA TTT GAA TCT TCA TCA TAC GAA GGA TAC TTT      657
145      His Asp Asn Lys Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe
146      110 115 120
147      CTA GCT TGT GAA AAA GAG AGA GAC CTT TTT AAA CTC ATT TTG AAA AAA      705
148      Leu Ala Cys Glu Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys
149      125 130 135 140
150      GAG GAT GAA TTG GGG GAT AGA TCT ATA ATG TTC ACT GTT CAA AAC GAA      753
151      Glu Asp Glu Leu Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu
152      145 150 155

```

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/479,862DATE: 12/20/2000
TIME: 22:20:59

INPUT SET: S36235.raw

153 GAC TAGCTATTAA AATTTTCATGC CGGGCGCAGT GGCTCACGCC TGTAATCCCA 806
154 Asp
155 GCCCTTTGGG AGGCTGAGGC GGGCAGATCA CCAGAGGTCA GGTGTTCAAG ACCAGCCTGA 866
156 CCAACATGGT GAAACCTCAT CTCTACTAAA AATACTAAAA ATTAGCTGAG TGTAGTGACG 926
157 CATGCCCTCA ATCCCAGCTA CTCAAGAGGC TGAGGCAGGA GAATCACTTG CACTCCGGAG 986
158 GTAGAGGTTG TGGTGAGCCG AGATTGCACC ATTGCGCTCT AGCCTGGGCA ACAACAGCAA 1046
159 AACTCCATCT CAAAAAATAA AATAAATAA TAAACAAATA AAAAATTCAT AATGTGAAAA 1106
160 AAAAAAAAAA AAAA 1120

161

162 (2) INFORMATION FOR SEQ ID NO: 3:

163

164 (i) SEQUENCE CHARACTERISTICS:

165 (A) LENGTH: 135 base pairs

166 (B) TYPE: nucleic acid

167 (C) STRANDEDNESS: double

168 (D) TOPOLOGY: linear

169

170 (ii) MOLECULE TYPE: Genomic DNA

171

172 (vi) ORIGINAL SOURCE:

173 (A) ORGANISM: human

174 (F) TISSUE TYPE: placenta

175

176 (ix) FEATURE:

177 (A) NAME/KEY: exon

178 (B) LOCATION: 1..135

179 (C) IDENTIFICATION METHOD: S

180

181 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

182

183 AA AAC CTG GAA TCA GAT TAC TTT GGC AAG CTT GAA TCT AAA TTA TCA 47

184 Glu Asn Leu Glu Ser Asp Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser

185 -5 1 5 10

186 GTC ATA AGA AAT TTG AAT GAC CAA GTT CTC TTC ATT GAC CAA GGA AAT 95

187 Val Ile Arg Asn Leu Asn Asp Gln Val Leu Phe Ile Asp Gln Gly Asn

188 15 20 25

189 CGG CCT CTA TTT GAA GAT ATG ACT GAT TCT GAC TGT AGA G 135

190 Arg Pro Leu Phe Glu Asp Met Thr Asp Ser Asp Cys Arg Asp

191 30 35 40

192

193 (2) INFORMATION FOR SEQ ID NO: 4:

194

195 (i) SEQUENCE CHARACTERISTICS:

196 (A) LENGTH: 134 base pairs

197 (B) TYPE: nucleic acid

198 (C) STRANDEDNESS: double

199 (D) TOPOLOGY: linear

200

201 (ii) MOLECULE TYPE: Genomic DNA

202

203 (vi) ORIGINAL SOURCE:

204 (A) ORGANISM: human

205 (F) TISSUE TYPE: placenta

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/479,862DATE: 12/20/2000
TIME: 22:20:59

INPUT SET: S36235.raw

206
207 (ix) FEATURE:
208 (A) NAME/KEY: exon
209 (B) LOCATION: 1..134
210 (C) IDENTIFICATION METHOD: S
211
212
213 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
214
215 AT AAT GCA CCC CGG ACC ATA TTT ATT ATA AGT ATG TAT AAA GAT AGC 47
216 Asp Asn Ala Pro Arg Thr Ile Phe Ile Ile Ser Met Tyr Lys Asp Ser
217 40 45 50 55
218 CAG CCT AGA GGT ATG GCT GTA ACT ATC TCT GTG AAG TGT GAG AAA ATT 95
219 Gln Pro Arg Gly Met Ala Val Thr Ile Ser Val Lys Cys Glu Lys Ile
220 60 65 70
221 TCA ACT CTC TCC TGT GAG AAC AAA ATT ATT TCC TTT AAG 134
222 Ser Thr Leu Ser Cys Glu Asn Lys Ile Ile Ser Phe Lys
223 80 85
224
225 (2) INFORMATION FOR SEQ ID NO: 5:
226
227 (i) SEQUENCE CHARACTERISTICS:
228 (A) LENGTH: 87 base pairs
229 (B) TYPE: nucleic acid
230 (C) STRANDEDNESS: double
231 (D) TOPOLOGY: linear
232
233 (ii) MOLECULE TYPE: Genomic DNA
234
235 (vi) ORIGINAL SOURCE:
236 (A) ORGANISM: human
237 (F) TISSUE TYPE: placenta
238
239 (ix) FEATURE:
240 (A) NAME/KEY: exon
241 (B) LOCATION: 1..87
242 (C) IDENTIFICATION METHOD: S
243
244 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
245
246 GAATAAAG ATG GCT GCT GAA CCA GTA GAA GAC AAT TGC ATC AAC TTT GTG 50
247 Met Ala Ala Glu Pro Val Glu Asp Asn Cys Ile Asn Phe Val
248 -35 -30 -25
249 GCA ATG AAA TTT ATT GAC AAT ACG CTT TAC TTT ATA G 87
250 Ala Met Lys Phe Ile Asp Asn Thr Leu Tyr Phe Ile Ala
251 -20 -15 -10
252
253 (2) INFORMATION FOR SEQ ID NO:6:
254
255 (i) SEQUENCE CHARACTERISTICS:
256 (A) LENGTH: 12 base pairs
257 (B) TYPE: nucleic acid
258 (C) STRANDEDNESS: double

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/479,862

DATE: 12/20/2000
TIME: 22:20:59

INPUT SET: S36235.raw

Line

Error

Original Text

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/09/479,862

DATE: 12/20/2000
TIME: 22:21:00

INPUT SET: S36235.raw

< < THERE ARE NO ITEMS MISSING > >

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/09/479,862DATE: 12/20/2000
TIME: 22:21:00

INPUT SET: S36235.raw

Line	Original Text	Corrected Text
99	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
102	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
105	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
108	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
179	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
210	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
242	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
270	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
295	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
365	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
410	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
512	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
682	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
727	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
806	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
809	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
812	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
815	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
818	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
821	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
824	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
827	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
830	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
833	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
836	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
839	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
1088	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
1091	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
1094	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
1097	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
1100	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
1103	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
1106	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
1109	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
1112	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
1115	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
1118	(C) IDENTIFICATION METHODS: S	(C) IDENTIFICATION METHOD: S
1121	(C) IDENTIFICATION METHODS: E	(C) IDENTIFICATION METHOD: E
1646	(15) INFORMATION FOR SEQ ID NO: 15:	(2) INFORMATION FOR SEQ ID NO: 15: